

Arcee's MergeKit: A Toolkit for Merging Large Language Models

Year: 2024 | Citations: 158 | Authors: Charles Goddard, Shamane Siriwardhana, Malikeh Ehghaghi

Abstract

The rapid growth of open-source language models provides the opportunity to merge model checkpoints, combining their parameters to improve performance and versatility. Advances in transfer learning have led to numerous task-specific models, which model merging can integrate into powerful multitask models without additional training. MergeKit is an open-source library designed to support this process with an efficient and extensible framework suitable for any hardware. It has facilitated the merging of thousands of models, contributing to some of the world's most powerful open-source model checkpoints. The library is accessible at: <https://github.com/arcee-ai/mergekit>.